

ニューラルネットワークによる手書き平仮名の識別—人間よりも人間らしいコンピューターへのステップ

近藤正紀

新潟医療福祉大学 医療情報管理学科

【背景・目的】人工知能、特に Deep Learning という言葉が一般化して久しいが、実用化、特に製品化についてはまだまだ数多の困難が予想される。実用化にのっての最大の難関は、人工知能の信頼性に対する人々の懸念である。1970年代にスタンフォード大学で開発された Mycin が医療現場に投入されなかった理由として、コンピューターが誤った場合の責任の所在が問題となったとされている。

自動車の自動運転でも、事故発生時の責任の所在を解決しない限り、製品化は困難と考えられているし、医療情報学においても、これらは『人工知能の人格性に関わる問題』として議論されている。

社会が人工知能の出した結論を受け入れるか否かは、結局のところ人工知能が誤る確率が、ヒトが誤る確率を大幅に下回っていることが重要であろうと考えられる。そこで本研究では、従前に行った手書き数字識別精度向上の研究を進展させ、手書き平仮名の識別を行う。

【方法】本研究では、データとして国立研究法人産業技術総合研究所が研究用途に限定して公開している ETL Character Database¹⁾を用いる。このデータベースは旧電子技術総合研究所が 1973 年から 1984 年にかけて文字認識研究用に収集した、手書きもしくは印刷の約 120 万の文字画像データを集積したもので、ETL-1 から ETL-9 の 9 種類にまとめられており、本研究では ETL-8 内の ETL8B2 から、平仮名のみを抜き出して使用する。

データ取得時の OCR シートを見ると、平仮名 75 種には小さい文字、即ち“っ”“ゃ”“ゅ”“ょ”が含まれているが、データ提供者がこれら大きい文字と小さい文字を区別して OCR シートに記入していない場合がある。先行研究²⁾ではこの点について明確には言及はないが、大きい文字と小さい文字を別のものとして扱っていると読み取れる。本研究では大きい文字と小さい文字を区別しないようデータを調整し、71 種の文字をニューラルネットワークで識別する。また、ETL8B2 内の 12000 枚の平仮名文字画像をランダムに並べ替え、訓練用 8625 枚、検証用 1875 枚、テスト用 1500 枚に分割し、訓練用と検証用で学習を行い、テスト用で成績（正解率）を求めることとする。ニューラルネットワークの構造は従前の研究と比較するため基本的な 3 層構造とするが、プログラムは Python 3.5 上の TensorFlow 1.2 で再記述する。

【結果】構築したニューラルネットワークの構造と正解率

は表 1 のとおりである。

表 1 ネットワークの構造と正解率

入力層	ノード数		正解率 (最良値)
	中間層	出力層	
4032 = 64 x 63	225	71	0.9103
	400		0.9102
	625		0.9074

【考察】筆者は MNIST³⁾データを用いて 3 層ニューラルネットワークでの学習実験において表 2 の結果を得ている。MNIST データとは、Yann LeCun らが 1998 年に公開した手書き数字のグレースケール画像データで、機械学習の実験で標準的に用いられているものである。

表 2 中間層のノード数(上段)と正解率(最良値)(下段)

30	50	100	300	500	1000
0.9498	0.9606	0.9798	0.9811	0.9832	0.9850

(入力層: 784 = 28 x 28, 出力層: 10)

この結果と比較すると、MNIST データの場合は中間層のノード数と正解率の間には何らかの相関の存在が示唆されるが、本研究でははっきりしない。原因の一つとして日本人の書いた平仮名がかなり均質であることが考えられるが、更なる解析が必要である。また、本研究のニューラルネットワークに MNIST データを読み込ませた結果として 0.9106 が得られており、正解率の低さはネットワークの構造やプログラムに由来するものではなことが判る。なお、試験的に構築した図 1 の convolutional neural network (CNN) と MNIST データでは正解率 0.9926 を得ている。

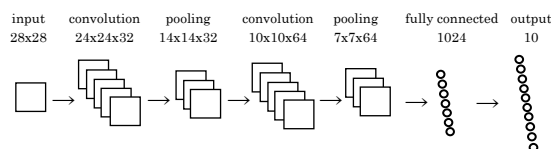


図 1 試作した convolutional neural network

【結論】日本人の書いた手書き平仮名の識別では、単純なニューラルネットワークでも 90%を超える正解率を得ることができた。今後は、ニューラルネットワークを改良し、Deep Learning によって漢字を含めて 99%を超える正解率を目指す。

【文献】

- 1) etlcdb, <http://etlcdb.db.aist.go.jp/>, 2017-8-21 アクセス
- 2) Charlie Tsai, Recognizing Handwritten Japanese Characters Using Deep Convolutional Neural Networks, http://cs231n.stanford.edu/reports/2016/pdfs/262_Report.pdf, 2016-6-22 アクセス
- 3) Yann LeCun, THE MNIST DATABASE of handwritten digits, <http://yann.lecun.com/exdb/mnist/>, 2017-8-21 アクセス